SF 298 MASTER COPY

KEEP THIS COPY FOR REPRODUCTION PURPOSES

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE: December 1996

3. REPORT TYPE AND DATES COVERED: Technical

4. TITLE AND SUBTITLE: Interval-Censored Type II Plan

5. FUNDING NUMBERS: DAAH04-96-1-0082

6. AUTHOR(S): S. Mudivarthy, M.B. Rao and R. Mitra

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)

Center for Multivariate Analysis
Department of Statistics
417 Thomas Building
Penn State University
University Park, PA 16802

8. PERFORMING ORGANIZATION REPORT NUMBER: 96-13

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER: ARO 35518.5-MA

11. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The Type II plan is used on many occasions in order to estimate the lifetime distribution of a product under investigation. Begin the process by choosing two positive integers $r \leq n$. Select a sample of n units of the product, set them to work, and observe the units continuously until r units fail. The objective is to estimate the lifetime distribution of the product using the data on r failure times.

We want to offer a modification of this plan in response to a past consultation problem. This problem arose from two diverse fields: one from engineering and the other from ornithology. In order to expedite observation of the r failure times, only periodic inspections were made. In such a plan, the exact failure of the r units will be unknown, but we will know how many units failed between each of the consecutive inspection times.

14. SUBJECT TERMS: Censored data, efficiency, interval-censored data, Type II plan

15. NUMBER IF PAGES: 6

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18
298-102

INTERVAL-CENSORED TYPE II PLAN

S. Mudivarthy, M.B. Rao, R. Mitra

Technical Report 96-13

December 1996

Center for Multivariate Analysis
417 Thomas Building
Penn State University
University Park, PA 16802

Research sponsored by the Army Research Office under Grant DAAHO4-96-1-0082. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

INTERVAL-CENSORED TYPE II PLAN

by

S. Mudivarthy, M.B. Rao[1]

Department of Statistics
North Dakota State University

R. Mitra

Moorhead State University

ABSTRACT

The Type II plan is used on many occasions in order to estimate the lifetime distribution of a product under investigation. Begin the process by choosing two positive integers $r \leq n$. Select a sample of n units of the product, set them to work, and observe the units continuously until r units fail. The objective is to estimate the lifetime distribution of the product using the data on r failure times.

We want to offer a modification of this plan in response to a past consultation problem. This problem arose from two diverse fields: one from engineering and the other from ornithology. In order to expedite observation of the r failure times, only periodic inspections were made. In such a plan, the exact failure of the r units will be unknown, but we will know how many units failed between each of the consecutive inspection times.

AMS Classification index: 62H30, 62H17.

Key Words and Phrases: Censored data, efficiency, interval-censored data, Type II plan.

[1]The research work of M.B. Rao is supported by Army Research Office Grant DAAH04-96-1-0082

INTERVAL-CENSORED TYPE II PLAN

Surekha Mudivarthy, M. Bharkara Rao, North Dakota State University
Rupa Mitra, Moorhead State University
M.B. Rao, North Dakota State University, Department of Statistics, P.O. Box 5575, Fargo, ND 58105

Key Words: Censored data, efficiency, interval-censored data, Type II plan

1. Introduction: The Type II plan is used on many occasions in order to estimate the lifetime distribution of a product under investigation. Begin the process by choosing two positive integers $r \leqq n$. Select a sample of n units of the product, set them to work, and observe the units continuously until r units fail. The objective is to estimate the lifetime distribution of the product using the data on r failure times.

We want to offer a modification of this plan in response to a past consultation problem. This problem arose from two diverse fields: one from engineering and the other from ornithology. In order to expedite observation of the r units failure times, only periodic inspections were made. In such a plan, the exact failure times of the r units will be unknown, but we will know how many units failed between each of the consecutive inspection times.

Formally, the inspection plan can be described as follows. Choose and fix a number $t_0 > 0$. Select a sample of n units and set them to work. Inspect the units at times $t_0, 2t_0, \ldots$ until r units fail. Let M denote the number of inspections needed. Let X_i = Number of units failed during the ith inspection

interval $\left((i-1)t_0, it_0\right]$,

i=1,2,3,.... The data consist of

$$M,\ X_1,\ X_2,\ \ldots,\ X_M.$$

These random variables satisfy the following conditions:

$$X_1 + X_2 + \ldots + X_{M-1} \leqq r - 1,$$

and

$$X_1 + X_2 + \ldots + X_M \geqq r.$$

We call this plan Interval-Censored Type II plan. The basic questions we were asked to address were:
1. Evaluate the loss of information in some meaningful way if one adopts the Interval-Censored Type II plan over the traditional Type II plan;
2. Provide some guidelines as to the choice of t_0.

In this paper, we attempt to answer these questions to the best of our ability. This is ongoing work and we want to report what we have achieved so far. Let T be the underlying lifetime variable associated with the product. We assume that T has an exponential distribution with unknown parameter $\theta > 0$. The probability density function of T is given by

$$f_\theta(x) = \theta e^{-\theta x},\ x > 0,\ \theta > 0.$$

We obtain the likelihood estimator of θ under the Interval-Censored Type II plan and compare its variance with the variance of the unbiased estimator of θ built on the maximum likelihood estimator of θ under the Type II plan. This comparison is done using extensive simulation studies. Some relevant distribution theory and asymptotics are presented here.

2. Some Distribution Theory: Let $T_1, T_2, \ldots, T_n$ be *iid* copies of T and $T_{(1)} < T_{(2)} < \ldots < T_{(n)}$ the corresponding order statistics. Under the Type II plan, the data consist of $T_{(1)},\ T_{(2)},\ \ldots,\ T_{(r)}$. The maximum likelihood estimator of θ is given by

$$\hat{\theta}_{r,n} = r \Big/ \left(T_{(1)} + T_{(2)} + \ldots + T_{(r)} + (n-r)T_{(r)}\right).$$

Under θ, $2r\theta / \hat{\theta}_{r,n}$ has a chi-squared distribution with $2r$ degree of freedom. The estimator $\hat{\theta}_{r,n}$ is biased, but the bias is correctable. More precisely,

$$E_\theta\left(\hat{\theta}_{r,n}\right) = \left(r/(r-1)\right)\theta \text{ for all } \theta > 0.$$

The unbiased estimator $\left((r-1)/r\right)\hat{\theta}_{r,n}$ has variance $\theta^2 / (r-2)$ provided $r > 2$. (Epstein and Sobel (1953).)

Let us focus on the Interval-Censored Type II plan. We need to determine the joint distribution of the data $M,\ X_1,\ X_2,\ \ldots,\ X_M$. We proceed in two stages. First, we obtain the distribution of M and then the conditional distribution of

$$X_1,\ X_2,\ \ldots,\ X_M \mid M = m.$$

Distribution of M: $\Pr_\theta(M=1) = \Pr_\theta(X_1 \geqq r)$

$= \Pr_\theta$ (at least r failures in the interval $(0, t_0]$)

$$= \sum_{x=r}^{n} \binom{n}{r} \left(1-e^{-\theta t_0}\right)^x \left(e^{-\theta t_0}\right)^{n-x}$$

$$= \int_0^{p_1} \frac{1}{B(r, n-r+1)} z^{r-1}(1-z)^{n-r}\, dz$$

$$= I_{p_1}(r, n-r+1),$$

where $p_1 = 1-e^{-\theta t_0}$ and $I_{p_1}(.,.)$ is the Incomplete Beta Function. (Abramowitz and Stegun (1965).) For $m \geqq 2$,

$\Pr_\theta(M=m)$

$= \Pr_\theta(X_1 + X_2 + \ldots + X_{m-1} \leqq r-1$ and $X_1 + X_2 + \ldots + X_m \leqq r)$

$= \sum_{s=0}^{r-1} \sum_{t=r-s}^{n-s} \Pr_\theta$ (s units fail in the interval $(0, (m-1)t_0]$ and t units fail in the interval $((m-1)t_0, mt_0])$

$$= \sum_{s=0}^{r-1} \sum_{t=r-s}^{n-s} \frac{n!}{s!\,t!\,(n-s-t)!} \left(1-e^{-\theta(m-1)t_0}\right)^s \left(e^{-\theta(m-1)t_0} - e^{-\theta m t_0}\right)^t \left(e^{-\theta m t_0}\right)^{n-s-t}$$

$$= \sum_{s=0}^{r-1} \frac{n!}{s!(n-s)!} \left(1-e^{-\theta(m-1)t_0}\right)^s \left(e^{-\theta(m-1)t_0}\right)^{n-s} \bullet \sum_{t=r-s}^{n-s} \frac{(n-s)!}{t!(n-s-t)!} \left(1-e^{-\theta t_0}\right)^t \left(e^{-\theta t_0}\right)^{n-s-t}$$

$$= \sum_{s=0}^{r-1} \frac{n!}{s!(n-s)!} p_{m-1}^s \left(1-p_{m-1}\right)^{n-s} I_{p_1}(r-s, n-r+1),$$

where $p_i = \left(1-e^{-\theta i t_0}\right)$, $i=1,2,\ldots$.

Conditional distribution of

$$X_1, X_2, \ldots, X_M | M = m:$$

Under $\theta > 0$,

$$\Pr_\theta(X_1 = x_1, X_2 = x_2, \ldots, X_m = x_m | M = m)$$

$$= \frac{n!}{x_1!\,x_2!\ldots x_m!\left(n-\sum_{i=1}^m x_i\right)!} \left(1-e^{-\theta t_0}\right)^{x_1} \left(e^{-\theta t_0} - e^{-2t_0\theta}\right)^{x_2} \ldots \left(e^{-(m-1)t_0\theta} - e^{-mt_0\theta}\right)^{x_m} \left(e^{-mt_0\theta}\right)^{\left(n-\sum_{i=1}^m x_i\right)},$$

for all $0 \leqq x_1, x_2, \ldots, x_m \leqq n$, $\sum_{i=1}^{m-1} x_i \leqq r-1$, and $\sum_{i=1}^{m} x_i \geqq r$. For the sake of simplicity, let

$$x_{m+1} = n - \sum_{i=1}^m x_i.$$

We now derive the maximum likelihood estimate of θ based on the data: number of inspections made and the number of units failed in each inspection interval, i.e.,

$$M = m, X_1 = x_1, X_2 = x_2, \ldots, X_m = x_m.$$

The likelihood L of the data is:

$$\mathrm{L} = \text{Constant}\left(1-\mathrm{e}^{-\theta t_0}\right)^{x_1} \left(e^{-\theta t_0} - e^{-2t_0\theta}\right)^{x_2} \ldots \left(e^{-(m-1)t_0\theta} - e^{-mt_0\theta}\right)^{x_m} \left(e^{-mt_0\theta}\right)^{x_{m+1}}$$

$$= \text{Constant}\left(1-\mathrm{e}^{-\theta t_0}\right)^{x_1+x_2+\ldots+x_m} \left(e^{-\theta t_0}\right)^{x_2+2x_3+3x_4+\ldots+mx_{m+1}}.$$

The log likelihood is given by:

$$\ln\ \mathrm{L} = \text{Constant} + \left(\sum_{i=1}^m x_i\right) \ln\left(1-e^{-\theta t_0}\right) - (\theta t_0)\left(\sum_{i=1}^m i x_{i+1}\right).$$

The derivative of the log likelihood is set equal to zero in order to obtain the maximum likelihood estimate. The following equations achieve the objective.

$$\frac{\partial}{\partial\theta}(\ln L) = \left(\sum_{i=1}^m x_i\right) \frac{t_0 e^{-\theta t_0}}{\left(1-e^{-\theta t_0}\right)} - t_0\left(\sum_{i=1}^m i x_{i+1}\right) = 0$$

$$\frac{e^{-\theta t_0}}{\left(1-e^{-\theta t_0}\right)} = \frac{\sum_{i=1}^m i x_{i+1}}{\sum_{i=1}^m x_i}$$

$$e^{-\theta t_0} = \frac{\sum_{i=1}^{m} i x_{i+1}}{\sum_{i=1}^{m} x_i + \sum_{i=1}^{m} i x_{i+1}}$$

$$\theta = -\frac{1}{t_0}\ln\left(\frac{\sum_{i=1}^{m} i x_{i+1}}{\sum_{i=1}^{m} x_i + \sum_{i=1}^{m} i x_{i+1}}\right)$$

$$= \frac{1}{t_0}\ln\left(1 + \frac{\sum_{i=1}^{m} x_i}{\sum_{i=1}^{m} i x_{i+1}}\right).$$

The maximum likelihood estimator of θ, upon replacing the data by the corresponding random variables, is given by

$$\hat{\theta} = \frac{1}{t_0}\ln\left(1 + \frac{\sum_{i=1}^{M} X_i}{\sum_{i=1}^{M} i X_{i+1}}\right).$$

The next objective is to obtain the asymptotic variance of $\hat{\theta}$. Rewrite the derivative of the log likelihood as

$$\frac{\partial}{\partial\theta}(\ln L) = t_0\left(\sum_{i=1}^{M} X_i\right)\left(\frac{1}{1-e^{-\theta t_0}} - 1\right)$$
$$-t_0\left(\sum_{i=1}^{m} i X_{i+1}\right),$$

from which we have

$$\frac{\partial^2}{\partial\theta^2}(\ln L) = -t_0^2\left(\sum_{i=1}^{M} X_i\right)\left(\frac{e^{-\theta t_0}}{\left(1-e^{-\theta t_0}\right)^2}\right).$$

The asymptotic variance of $\hat{\theta}$ is given by

$$\mathrm{Asyvar}_\theta(\hat{\theta}) = \frac{1}{E_\theta\left(-\frac{\partial^2}{\partial\theta^2}(\ln L)\right)}$$

$$= \frac{e^{\theta t_0}\left(1-e^{-\theta t_0}\right)^2}{t_0^2 E_\theta\left(\sum_{i=1}^{M} X_i\right)}.$$

The formula for the asymptotic variance simplifies to evaluating successfully $E\left(\sum_{i=1}^{M} X_i\right)$. We will evaluate the expectation using the conditional expectation argument. Note that

$$E_\theta\left(\sum_{i=1}^{M} X_i\right) = E\left(E\left(\sum_{i=1}^{M} X_i \middle| M\right)\right)$$

$$= \sum_{m \geq 1} E\left(\sum_{i=1}^{m} X_i \middle| M=m\right)\Pr_\theta(M=m).$$

The critical step is the evaluation of the conditional expectation:

$$E\left(\sum_{i=1}^{M} X_i \middle| M=m\right)$$

$$= \Sigma\left(x_1+x_2+\ldots+x_m\right)\frac{n!}{x_1!\,x_2!\ldots x_m!\,x_{m+1}!}$$
$$\left(1-e^{-\theta t_0}\right)^{x_1}$$
$$\left(e^{-\theta t_0} - e^{-2t_0\theta}\right)^{x_2}\ldots\left(e^{-(m-1)t_0\theta} - e^{-mt_0\theta}\right)^{x_m}$$
$$\left(e^{-mt_0\theta}\right)^{x_{m+1}},$$

where the summation is taken over all integers

$$0 \leqq x_1, x_2, \ldots, x_m, x_{m+1} \leqq n,$$
$$x_1, x_2, \ldots, x_{m+1} \leqq r-1,$$
$$x_1, x_2, \ldots, x_m \geqq r,$$

and

$$x_1, x_2, \ldots, x_{m+1} = n.$$

Writing $x_1, x_2, \ldots, x_{m+1} = s$ and $x_m = t$, we can rewrite the conditional expectation as

$$E\left(\sum_{i=1}^{m} X_i \middle| M=m\right)$$

$$= \sum_{s=0}^{r-1}\sum_{t=r-s}^{n-s}(s+t)\frac{n!}{s!\,t!(n-s-t)!}$$
$$\sum_{\substack{x_1,x_2,\ldots,x_{m-1}\geq 0 \\ x_1+x_2+\ldots+x_{m-1}=s}} \frac{s!}{x_1!\,x_2!\ldots x_{m-1}!}$$
$$\left(1-e^{-\theta t_0}\right)^{x_1}\left(e^{-\theta t_0} - e^{-2t_0\theta}\right)^{x_2}\ldots$$
$$\left(e^{-(m-1)t_0\theta} - e^{-mt_0\theta}\right)^{x_m}\left(e^{-mt_0\theta}\right)^{x_{m+1}}$$

$$= \sum_{s=0}^{r-1}\sum_{t=r-s}^{n-s}(s+t)\frac{n!}{s!\,t!(n-s-t)!}$$

$$\left[\left(1-e^{-\theta t_0}\right)+\left(e^{-\theta t_0}-e^{-2t_0\theta}\right)+\ldots +\left(e^{-(m-2)t_0\theta}-e^{-(m-1)t_0\theta}\right)\right]^s$$

$$\left(e^{-(m-1)t_0\theta}-e^{-mt_0\theta}\right)^t\left(e^{-mt_0\theta}\right)^{n-s-t}$$

$$=\sum_{s=0}^{r-1}\sum_{t=r-s}^{n-s}(s+t)\frac{n!}{s!t!(n-s-t)!}$$

$$\left(1-e^{-(m-1)t_0\theta}\right)^s\left(e^{-(m-1)t_0\theta}-e^{-mt_0\theta}\right)^t$$

$$\left(e^{-mt_0\theta}\right)^{n-s-t}$$

$$=\sum_{s=0}^{r-1}\sum_{t=r-s}^{n-s}(s)\frac{n!}{s!t!(n-s-t)!}\left(1-e^{-(m-1)t_0\theta}\right)^s$$

$$\left(e^{-(m-1)t_0\theta}-e^{-mt_0\theta}\right)^t\left(e^{-mt_0\theta}\right)^{n-s-t}$$

$$+\sum_{s=0}^{r-1}\sum_{t=r-s}^{n-s}(t)\frac{n!}{s!t!(n-s-t)!}\left(1-e^{-(m-1)t_0\theta}\right)^s$$

$$\left(e^{-(m-1)t_0\theta}-e^{-mt_0\theta}\right)^t\left(e^{-mt_0\theta}\right)^{n-s-t}$$

$$=n\left(1-e^{-(m-1)t_0\theta}\right)\sum_{s=0}^{r-1}\frac{(n-1)!}{(s-1)!(n-s)!}$$

$$\left(1-e^{-(m-1)t_0\theta}\right)^{s-1}\left(e^{-(m-1)t_0\theta}\right)^{n-s}$$

$$I_{p_1}(r-s,n-r+1)$$

$$+(n-s)\left(1-e^{-\theta t_0}\right)\sum_{t=r-s}^{n-s}\frac{(n-s-1)!}{(t-1)!(n-s-t)!}\bullet$$

$$\left(1-e^{-\theta t_0}\right)^{t-1}\left(e^{-\theta t_0}\right)^{n-s-t}.$$

$$\left(1-I_{p_{m-1}}(r,n-r+1)\right).$$

3. Simulations: The mean square error of the maximum likelihood estimator of θ under interval-censored Type II plan is intractable. We evaluated the mean square error of the maximum likelihood estimator empirically by mounting Monte Carlo studies. The inputs are:

n = sample size (10,20,30,40);
r = no. of failures allowed to be observed;
θ = 0.1, 1, 10;
t_0 = length of the time interval.

For each choice of r, θ, and t_0; 5,000 samples each of size n were drawn, maximum likelihood estimate of θ computed, and the empirical mean square error evaluated. The objective is to compare the mean square with the variance of the unbiased estimator of θ based on the likelihood of the data under the continuous inspection Type II plan for the same choice of r, θ, and t_0. The efficiency of the interval-censored Type II plan is computed by computing the ratio,

$$\frac{\text{variance under continuous inspection plan}}{\text{mean square error under interval - censored plan}}.$$

A part of our computational effort is presented below.

Sample size n=20

r	t_0	θ	Efficiency
10	0.5	1	0.9878
10	0.6	1	1.0551
10	0.7	1	1.1103
10	0.8	1	1.1252
10	0.9	1	1.0882
10	1.0	1	1.0904
15	0.5	1	0.9537
15	0.6	1	0.9306
15	0.7	1	0.9502
15	0.8	1	0.9298
15	0.9	1	0.9116
15	1.0	1	0.8532

4. Conclusion:

Even for a moderate sample size like $n = 20$ and for a moderate value of $r = 10$, the interval-censored Type II plan is as good as the Type II plan, if not better. More work is needed to make an overall recommendation.

Acknowledgements: The authors are grateful to Mr. Jerry Stockrahm for a careful reading of the paper.

***

References

1. Abramowitz, M., and Stegun, A.I. (1965). Handbook of Mathematical Functions. Dover Publications, Inc., New York.

2. Epstein, B., and Sobel, M. (1953). "Life Testing", J. Amer. Stat. Assoc., 48, 486-502.

INTERVAL-CENSORED TYPE II PLAN

Surekha Mudivarthy, M. Bharkara Rao, North Dakota State University
Rupa Mitra, Moorhead State University
M.B. Rao, North Dakota State University, Department of Statistics, P.O. Box 5575, Fargo, ND 58105

Key Words: Censored data, efficiency, interval-censored data, Type II plan

1. Introduction: The Type II plan is used on many occasions in order to estimate the lifetime distribution of a product under investigation. Begin the process by choosing two positive integers $r \leqq n$. Select a sample of n units of the product, set them to work, and observe the units continuously until r units fail. The objective is to estimate the lifetime distribution of the product using the data on r failure times.

We want to offer a modification of this plan in response to a past consultation problem. This problem arose from two diverse fields: one from engineering and the other from ornithology. In order to expedite observation of the r units failure times, only periodic inspections were made. In such a plan, the exact failure times of the r units will be unknown, but we will know how many units failed between each of the consecutive inspection times.

Formally, the inspection plan can be described as follows. Choose and fix a number $t_0 > 0$. Select a sample of n units and set them to work. Inspect the units at times $t_0, 2t_0, \ldots$ until r units fail. Let M denote the number of inspections needed. Let X_i = Number of units failed during the ith inspection interval $((i-1)t_0, it_0]$,

i=1,2,3,.... The data consist of

$$M,\ X_1,\ X_2,\ \ldots,\ X_M.$$

These random variables satisfy the following conditions:

$$X_1 + X_2 + \ldots + X_{M-1} \leqq r - 1,$$

and

$$X_1 + X_2 + \ldots + X_M \geqq r.$$

We call this plan Interval-Censored Type II plan. The basic questions we were asked to address were:

1. Evaluate the loss of information in some meaningful way if one adopts the Interval-Censored Type II plan over the traditional Type II plan;
2. Provide some guidelines as to the choice of t_0.

In this paper, we attempt to answer these questions to the best of our ability. This is ongoing work and we want to report what we have achieved so far. Let T be the underlying lifetime variable associated with the product. We assume that T has an exponential distribution with unknown parameter $\theta > 0$. The probability density function of T is given by

$$f_\theta(x) = \theta e^{-\theta x},\ x > 0,\ \theta > 0.$$

We obtain the likelihood estimator of θ under the Interval-Censored Type II plan and compare its variance with the variance of the unbiased estimator of θ built on the maximum likelihood estimator of θ under the Type II plan. This comparison is done using extensive simulation studies. Some relevant distribution theory and asymptotics are presented here.

2. Some Distribution Theory: Let $T_1, T_2, \ldots, T_n$ be *iid* copies of T and $T_{(1)} < T_{(2)} < \ldots < T_{(n)}$ the corresponding order statistics. Under the Type II plan, the data consist of $T_{(1)},\ T_{(2)},\ \ldots,\ T_{(r)}$. The maximum likelihood estimator of θ is given by

$$\hat{\theta}_{r,n} = r \Big/ \left(T_{(1)} + T_{(2)} + \ldots + T_{(r)} + (n-r)T_{(r)}\right).$$

Under θ, $2r\theta / \hat{\theta}_{r,n}$ has a chi-squared distribution with $2r$ degree of freedom. The estimator $\hat{\theta}_{r,n}$ is biased, but the bias is correctable. More precisely,

$$E_\theta\left(\hat{\theta}_{r,n}\right) = \left(r/(r-1)\right)\theta \text{ for all } \theta > 0.$$

The unbiased estimator $((r-1)/r)\hat{\theta}_{r,n}$ has variance $\theta^2 / (r-2)$ provided $r > 2$. (Epstein and Sobel (1953).)

Let us focus on the Interval-Censored Type II plan. We need to determine the joint distribution of the data $M,\ X_1,\ X_2,\ \ldots,\ X_M$. We proceed in two stages. First, we obtain the distribution of M and then the conditional distribution of

$$X_1,\ X_2,\ \ldots,\ X_M | M = m.$$

Distribution of M: $\Pr_\theta(M=1)=\Pr_\theta(X_1 \geqq r)$

$= \Pr_\theta$ (at least r failures in the interval $(0,t_0]$)

$$= \sum_{x=r}^{n}\binom{n}{r}\left(1-e^{-\theta t_0}\right)^x\left(e^{-\theta t_0}\right)^{n-x}$$

$$= \int_0^{p_1}\frac{1}{B(r,n-r+1)}z^{r-1}(1-z)^{n-r}\,dz$$

$$= I_{p_1}(r,n-r+1),$$

where $p_1 = 1-e^{-\theta t_0}$ and $I_{p_1}(.,.)$ is the Incomplete Beta Function. (Abramowitz and Stegun (1965).) For $m \geqq 2$,

$\Pr_\theta(M=m)$

$= \Pr_\theta(X_1+X_2+\ldots+X_{m-1} \leqq r-1$ and $X_1+X_2+\ldots+X_m \leqq r)$

$= \sum_{s=0}^{r-1}\sum_{t=r-s}^{n-s}\Pr_\theta$ (s units fail in the interval $(0,(m-1)t_0]$ and t units fail in the interval $((m-1)t_0, mt_0])$

$$= \sum_{s=0}^{r-1}\sum_{t=r-s}^{n-s}\frac{n!}{s!\,t!\,(n-s-t)!}\left(1-e^{-\theta(m-1)t_0}\right)^s \left(e^{-\theta(m-1)t_0}-e^{-\theta m t_0}\right)^t\left(e^{-\theta m t_0}\right)^{n-s-t}$$

$$= \sum_{s=0}^{r-1}\frac{n!}{s!(n-s)!}\left(1-e^{-\theta(m-1)t_0}\right)^s \left(e^{-\theta(m-1)t_0}\right)^{n-s} \bullet \sum_{t=r-s}^{n-s}\frac{(n-s)!}{t!(n-s-t)!}\left(1-e^{-\theta t_0}\right)^t\left(e^{-\theta t_0}\right)^{n-s-t}$$

$$= \sum_{s=0}^{r-1}\frac{n!}{s!(n-s)!}p_{m-1}^s\left(1-p_{m-1}\right)^{n-s} I_{p_1}(r-s,n-r+1),$$

where $p_i = \left(1-e^{-\theta i t_0}\right)$, i=1,2,... .

Conditional distribution of
$$X_1,X_2,\ldots,X_M \mid M=m:$$

Under $\theta>0$,

$\Pr_\theta(X_1=x_1,X_2=x_2,\ldots,X_m=x_m \mid M=m)$

$$= \frac{n!}{x_1!x_2!\ldots x_m!\left(n-\sum_{i=1}^m x_i\right)!}\left(1-e^{-\theta t_0}\right)^{x_1} \left(e^{-\theta t_0}-e^{-2t_0\theta}\right)^{x_2}\ldots\left(e^{-(m-1)t_0\theta}-e^{-mt_0\theta}\right)^{x_m} \left(e^{-mt_0\theta}\right)^{\left(n-\sum_{i=1}^m x_i\right)},$$

for all $0 \leqq x_1, x_2, \ldots, x_m \leqq n$, $\sum_{i=1}^{m-1} x_i \leqq r-1$, and $\sum_{i=1}^{m} x_i \geqq r$. For the sake of simplicity, let

$$x_{m+1} = n-\sum_{i=1}^m x_i .$$

We now derive the maximum likelihood estimate of θ based on the data: number of inspections made and the number of units failed in each inspection interval, i.e.,

$$M=m, X_1=x_1, X_2=x_2,\ldots,X_m=x_m.$$

The likelihood L of the data is:

$$L = \text{Constant}\left(1-e^{-\theta t_0}\right)^{x_1}\left(e^{-\theta t_0}-e^{-2t_0\theta}\right)^{x_2}\ldots \left(e^{-(m-1)t_0\theta}-e^{-mt_0\theta}\right)^{x_m}\left(e^{-mt_0\theta}\right)^{x_{m+1}}$$

$$= \text{Constant}\left(1-e^{-\theta t_0}\right)^{x_1+x_2+\ldots+x_m} \left(e^{-\theta t_0}\right)^{x_2+2x_3+3x_4+\ldots+mx_{m+1}}.$$

The log likelihood is given by:

$$\ln L = \text{Constant} + \left(\sum_{i=1}^m x_i\right)\ln\left(1-e^{-\theta t_0}\right) - \left(\theta t_0\right)\left(\sum_{i=1}^m i x_{i+1}\right).$$

The derivative of the log likelihood is set equal to zero in order to obtain the maximum likelihood estimate. The following equations achieve the objective.

$$\frac{\partial}{\partial\theta}(\ln L) = \left(\sum_{i=1}^m x_i\right)\frac{t_0 e^{-\theta t_0}}{\left(1-e^{-\theta t_0}\right)} - t_0\left(\sum_{i=1}^m i x_{i+1}\right) = 0$$

$$\frac{e^{-\theta t_0}}{\left(1-e^{-\theta t_0}\right)} = \frac{\sum_{i=1}^m i x_{i+1}}{\sum_{i=1}^m x_i}$$

$$e^{-\theta t_0} = \frac{\sum_{i=1}^{m} i x_{i+1}}{\sum_{i=1}^{m} x_i + \sum_{i=1}^{m} i x_{i+1}}$$

$$\theta = -\frac{1}{t_0}\ln\left(\frac{\sum_{i=1}^{m} i x_{i+1}}{\sum_{i=1}^{m} x_i + \sum_{i=1}^{m} i x_{i+1}}\right)$$

$$= \frac{1}{t_0}\ln\left(1 + \frac{\sum_{i=1}^{m} x_i}{\sum_{i=1}^{m} i x_{i+1}}\right).$$

The maximum likelihood estimator of θ, upon replacing the data by the corresponding random variables, is given by

$$\hat{\theta} = \frac{1}{t_0}\ln\left(1 + \frac{\sum_{i=1}^{M} X_i}{\sum_{i=1}^{M} i X_{i+1}}\right).$$

The next objective is to obtain the asymptotic variance of $\hat{\theta}$. Rewrite the derivative of the log likelihood as

$$\frac{\partial}{\partial\theta}(\ln L) = t_0\left(\sum_{i=1}^{M} X_i\right)\left(\frac{1}{1-e^{-\theta t_0}} - 1\right)$$
$$-t_0\left(\sum_{i=1}^{m} i X_{i+1}\right),$$

from which we have

$$\frac{\partial^2}{\partial\theta^2}(\ln L) = -t_0^2\left(\sum_{i=1}^{M} X_i\right)\left(\frac{e^{-\theta t_0}}{\left(1-e^{-\theta t_0}\right)^2}\right).$$

The asymptotic variance of $\hat{\theta}$ is given by

$$\mathrm{Asyvar}_\theta(\hat{\theta}) = \frac{1}{E_\theta\left(-\frac{\partial^2}{\partial\theta^2}(\ln L)\right)}$$

$$= \frac{e^{\theta t_0}\left(1-e^{-\theta t_0}\right)^2}{t_0^2 E_\theta\left(\sum_{i=1}^{M} X_i\right)}.$$

The formula for the asymptotic variance simplifies to evaluating successfully $E\left(\sum_{i=1}^{M} X_i\right)$. We will evaluate the expectation using the conditional expectation argument. Note that

$$E_\theta\left(\sum_{i=1}^{M} X_i\right) = E\left(E\left(\sum_{i=1}^{M} X_i \middle| M\right)\right)$$

$$= \sum_{m \geq 1} E\left(\sum_{i=1}^{m} X_i \middle| M = m\right)\Pr_\theta(M = m).$$

The critical step is the evaluation of the conditional expectation:

$$E\left(\sum_{i=1}^{M} X_i \middle| M = m\right)$$

$$= \Sigma\left(x_1 + x_2 + \ldots + x_m\right)\frac{n!}{x_1!\,x_2!\ldots x_m!\,x_{m+1}!}$$
$$\left(1-e^{-\theta t_0}\right)^{x_1}$$
$$\left(e^{-\theta t_0} - e^{-2t_0\theta}\right)^{x_2}\ldots\left(e^{-(m-1)t_0\theta} - e^{-mt_0\theta}\right)^{x_m}$$
$$\left(e^{-mt_0\theta}\right)^{x_{m+1}},$$

where the summation is taken over all integers

$$0 \leqq x_1, x_2, \ldots, x_m, x_{m+1} \leqq n,$$
$$x_1, x_2, \ldots, x_{m+1} \leqq r-1,$$
$$x_1, x_2, \ldots, x_m \geqq r,$$

and

$$x_1, x_2, \ldots, x_{m+1} = n.$$

Writing $x_1, x_2, \ldots, x_{m+1} = s$ and $x_m = t$, we can rewrite the conditional expectation as

$$E\left(\sum_{i=1}^{m} X_i \middle| M = m\right)$$

$$= \sum_{s=0}^{r-1}\sum_{t=r-s}^{n-s}(s+t)\frac{n!}{s!\,t!\,(n-s-t)!}$$

$$\sum_{\substack{x_1, x_2, \ldots, x_{m-1} \geq 0 \\ x_1 + x_2 + \ldots + x_{m-1} = s}} \frac{s!}{x_1!\,x_2!\ldots x_{m-1}!}$$

$$\left(1-e^{-\theta t_0}\right)^{x_1}\left(e^{-\theta t_0} - e^{-2t_0\theta}\right)^{x_2}\ldots$$

$$\left(e^{-(m-1)t_0\theta} - e^{-mt_0\theta}\right)^{x_m}\left(e^{-mt_0\theta}\right)^{x_{m+1}}$$

$$= \sum_{s=0}^{r-1}\sum_{t=r-s}^{n-s}(s+t)\frac{n!}{s!\,t!\,(n-s-t)!}$$

$$\left[\left(1-e^{-\theta t_0}\right)+\left(e^{-\theta t_0}-e^{-2t_0\theta}\right)+\ldots +\left(e^{-(m-2)t_0\theta}-e^{-(m-1)t_0\theta}\right)\right]^s$$

$$\left(e^{-(m-1)t_0\theta}-e^{-mt_0\theta}\right)^t\left(e^{-mt_0\theta}\right)^{n-s-t}$$

$$=\sum_{s=0}^{r-1}\sum_{t=r-s}^{n-s}(s+t)\frac{n!}{s!t!(n-s-t)!}\left(1-e^{-(m-1)t_0\theta}\right)^s\left(e^{-(m-1)t_0\theta}-e^{-mt_0\theta}\right)^t\left(e^{-mt_0\theta}\right)^{n-s-t}$$

$$=\sum_{s=0}^{r-1}\sum_{t=r-s}^{n-s}(s)\frac{n!}{s!t!(n-s-t)!}\left(1-e^{-(m-1)t_0\theta}\right)^s\left(e^{-(m-1)t_0\theta}-e^{-mt_0\theta}\right)^t\left(e^{-mt_0\theta}\right)^{n-s-t}$$

$$+\sum_{s=0}^{r-1}\sum_{t=r-s}^{n-s}(t)\frac{n!}{s!t!(n-s-t)!}\left(1-e^{-(m-1)t_0\theta}\right)^s\left(e^{-(m-1)t_0\theta}-e^{-mt_0\theta}\right)^t\left(e^{-mt_0\theta}\right)^{n-s-t}$$

$$=n\left(1-e^{-(m-1)t_0\theta}\right)\sum_{s=0}^{r-1}\frac{(n-1)!}{(s-1)!(n-s)!}\cdot\left(1-e^{-(m-1)t_0\theta}\right)^{s-1}\left(e^{-(m-1)t_0\theta}\right)^{n-s} I_{p_1}(r-s,n-r+1)$$

$$+(n-s)\left(1-e^{-\theta t_0}\right)\sum_{t=r-s}^{n-s}\frac{(n-s-1)!}{(t-1)!(n-s-t)!}\bullet\left(1-e^{-\theta t_0}\right)^{t-1}\left(e^{-\theta t_0}\right)^{n-s-t}\left(1-I_{p_{m-1}}(r,n-r+1)\right).$$

3. Simulations: The mean square error of the maximum likelihood estimator of θ under interval-censored Type II plan is intractable. We evaluated the mean square error of the maximum likelihood estimator empirically by mounting Monte Carlo studies. The inputs are:

n = sample size (10,20,30,40);
r = no. of failures allowed to be observed;
θ = 0.1, 1, 10;
t_0 = length of the time interval.

For each choice of r, θ, and t_0, 5,000 samples each of size n were drawn, maximum likelihood estimate of θ computed, and the empirical mean square error evaluated. The objective is to compare the mean square with the variance of the unbiased estimator of θ based on the likelihood of the data under the continuous inspection Type II plan for the same choice of r, θ, and t_0. The efficiency of the interval-censored Type II plan is computed by computing the ratio,

$$\frac{\text{variance under continuous inspection plan}}{\text{mean square error under interval-censored plan}}.$$

A part of our computational effort is presented below.

Sample size n=20

r	t_0	θ	Efficiency
10	0.5	1	0.9878
10	0.6	1	1.0551
10	0.7	1	1.1103
10	0.8	1	1.1252
10	0.9	1	1.0882
10	1.0	1	1.0904
15	0.5	1	0.9537
15	0.6	1	0.9306
15	0.7	1	0.9502
15	0.8	1	0.9298
15	0.9	1	0.9116
15	1.0	1	0.8532

4. Conclusion:

Even for a moderate sample size like $n = 20$ and for a moderate value of $r = 10$, the interval-censored Type II plan is as good as the Type II plan, if not better. More work is needed to make an overall recommendation.

Acknowledgements: The authors are grateful to Mr. Jerry Stockrahm for a careful reading of the paper.

***

References

1. Abramowitz, M., and Stegun, A.I. (1965). Handbook of Mathematical Functions. Dover Publications, Inc., New York.

2. Epstein, B., and Sobel, M. (1953). "Life Testing", J. Amer. Stat. Assoc., 48, 486-502.